



Data Distribution Committee

Chris Miller (Michigan)
Benjamin Weaver (NYU)



What is the DDC?



- The DESI Construction Project will provide a **limited** set of **basic** data access tools.
 - [DESI-doc-908](#)
- The DESI Collaboration provides the rest.
- The DDC exists to identify what needs to be provided.
 - [Our Charge in full](#)

What Are We Doing Right Now?



- Speaking to you, of course.
- Collect Use Cases.
- Convert Use Cases into Requirements.
- Write a Report.
- Issue RFPs as needed.
- Think about data release policy.



Use Case: Spectrum Quality



- I want to access/analyze a DESI spectrum.
 - View pipeline summary statistics and fitted spectrum.
 - View simple Q & A measures.
 - Account for a large dynamic range of spectral types, wavelength coverage, and S/N.
 - Keep up-to-date with most recent data and reductions (“last night”).

Use Case: Intelligent Software



- I want to run science code that knows about DESI data.
 - “I don’t want to hunt for files.”
 - Importable library that knows how to obtain the data.
 - Be able to recreate the data structure locally at various levels of completeness.
 - No (or small) difference when running code locally or at NERSC.

Use Case: Backtracking Data



- I want to view the all the data for a spectrum/target/brick.
- Connect target DB to imaging and spectrum DBs.
 - Easy!
- Connect target **files** to imaging and spectrum **files**.
 - Not quite so easy! This requires careful thought.
 - Pass row numbers in tables?

Use Case: Backtracking Bricks



- I have an arbitrary RA, Dec and need to find the corresponding brick name.
- In theory, this involves formatting the RA, Dec properly, **if one also knows the brick size.**
- Will brick size be finalized? Who decides that? If not, how does someone find the current brick size?

Defining Interfaces



- The use cases get converted into requirements on interfaces.
- Many existing examples from other projects, but would be nice to have more.
- Core library → API → WebApp.

Data Release Policy



- Internal data releases eventually become external.
- Lag between pipeline results and catalog-level results (e.g. LSS catalogs).
 - Maybe releases can be phased to take this into account.

Links and Documents



- DDC on [Trac](#).
- The raw text for the [DataDistDoc](#) Report.
- View the draft Report at [DESI-doc-1048](#).



Conclusion



- We still want to hear about use cases and interfaces from other projects.
 - If only there were some DES people who also happened to work for DESI...
- If you have questions email `desi-data@desi.lbl.gov`.

